\ufp

NASA TECHNICAL TRANSLATION NASA TT F-13,871

MATHEMATICAL PROBLEMS OF COMPRESSING LARGE MASSES OF MEASUREMENT INFORMATION

A.S. Golubkov and A.V. Sharonov

N72-14175 (NASA-TT-F-13871) MATHEMATICAL PROBLEMS OF COMPRESSING LARGE MASSES OF MEASUREMENT INFORMATION A.S. Golubkov, et al (Scripta Technica, Inc.) Oct. 1971 5 p CSCL 09B

Unclas 12023 G3/07

13871

Translation of "Matematicheskiye voprosy szhatiya bolshikh massivov izmeritel'noy informatsii"
IN: Chetverty simpozium po probleme izbytochnosti v informationnykh sistemakh (Fourth Symposium on the Problem of Redundancy in Information Systems), Leningrad, USSR Academy of ~~XXXXX~~ Sciences, 1970, pp 123-129

Chetvertyi ...

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546 OCTOBER 1971

MATHEMATICAL PROBLEMS OF COMPRESSING LARGE MASSES OF MEASUREMENT INFORMATION /123*

A.S. Golubkov and A.V. Sharonov

A study is made of a method of compressing large masses of information that utilizes the properties of Markov chains with a countable phase space.

Work on the construction of new technological models requires, in addition to preliminary studies and calculations and the development of the models themselves, very complex and extensive laboratory, bench, and field experiments. /124

The information regarding the object in question can be obtained by measurement of the parameters of the processes that take place in it. Here, the large number of these parameters, the briefness of the processes, etc. make it necessary to develop efficient control methods.

Among the foremost of these methods is the compression of the information obtained on-board of the object.

In this connection, we consider a method that, under certain restrictions on the processes actually taking place, allows a compression of the measurement information.

We shall confine ourselves to Gaussian processes $\xi(t)$ with independent increments that, with probability 1, have continuous trajectories. We know that such processes belong to the class of diffusion processes [2].

Suppose also that we have a priori information regarding the range of values of $\xi(t)$, which characterizes the normal operation of the object. Then, the solution of the problem of compression of the measurement information can be represented as the result of transmission of only those values of $\xi(t)$ that get beyond the boundary of the admissible range of values.

Consider the space $\mathcal{X}$ consisting of the functions $\xi(t)$, for $t \geq 0$, that assume values in the 2-dimensional vector space R.

*Numbers in the margin indicate pagination in the foreign text.

Let G denote a region characterizing the normal operation of the object. Let τ denote the instant the trajectory of $\xi(t)$ first gets out of that region. /125

We shall be interested in the probability $P_\xi(\tau < \infty)$ that the trajectory beginning at an arbitrary point $\xi \in G$ will get out of G. Of course, there are infinitely many variants of the flow of process $\xi(t)$. However, what we are interested in is not the process itself but its increments in the course of a brief interval of time $(t_0, t_0 + \Delta t)$ at those points at which it is close to the boundary of the region G.

Then, instead of $\xi(t) - \xi(t_0)$, we can consider the linear function $\xi'(t_0)(t-t_0)$, which is determined by the single number $\xi'(t_0)$. This number completely determines the behavior of $\xi(t)$ in an arbitrarily small neighborhood of t_0. It is called the infinitesimal characteristic of our process at the instant t_0.

Here, it is important that it be possible to restore the process as a whole when we know such characteristics for all t. For processes satisfying the requirements imposed above, it has been shown that Laplacian operator, which describes the behavior of the process close to the given point is, up to a constant factor, such a characteristic.

Suppose that the region G is bounded by a smooth plane curve and that, at the instant τ at which the trajectory of the process $\xi(t)$ first gets out of G, the image point of the process is on the boundary L. The probability that $\xi(t)$ belongs to a definite section of the boundary L is a function of the initial point $\xi \in G$. Let us denote this function by $f(\xi)$. We note that /126

$$f(\xi) = M_\xi \varphi(\xi(\tau)) \tag{1}$$

where φ is the function defined on L by

$$\varphi(y) = \begin{cases} 1 \text{ for } y \in \Gamma, \\ 0 \text{ for } y \notin \Gamma. \end{cases}$$

In this connection, it is shown that the function $f(\xi)$ is a harmonic function in G, that is, that it satisfies Laplace's equation [1]

$$\Delta f = 0 \ .$$

Thus, the determination of the probality that the trajectory will get beyond the boundary of the region G is reduced to solving Dirichlet's problem. By virtue of the maximum principle, this solution is unique. For example, let us consider the instant τ at which a trajectory beginning in the upper half-plane gets onto the t-axis. Let (t, ξ_1) denote a point in the upper half-plane. Then the solution of this problem for a function $f(\xi)$ that is harmonic in the upper half-plane and assumes given values $\varphi(y)$ on the t-axis is given by the formula

$$f(\xi) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\xi_1 \varphi(y)\, dy}{\xi_1^2 + (y-t)^2} \tag{2}$$

By using the fact that a conformal mapping maps harmonic functions into harmonic functions, we can use formula (2) to solve the Dirichlet problem for various regions in the plane.

Consequently, if we fix the instants t and use the fact that the processes under consideration are Markov processes, that is, that their behavior after the instant t obeys the same rule as if the initial instant were t = 0, we can, for each such point, calculate the probability that a trajectory beginning at the point will get beyond the boundary of the region G. /127

In particular, the values of processes obtained over the course of intervals of time different from the period of interrogation of the sensors are such points.

With this approach to the problem of compression of information, we must begin the transmission at the instant at which the probability of the trajectory's getting out becomes sufficiently high. In other words, we consider the sequence of measurements

$$\xi(t_0), \xi(t_1), \ldots \xi(t_n)$$

from which we need to choose the "best measurement" in the sense that the probability that a trajectory beginning at the point $(t_i, \xi(t_i))$ (for i = 0, 1, ..., n) will get out of the region G must be as great as possible.[1] Also, a return to measurements already considered is inadmissible. Since the sequence of measurements

$$\xi(t_0), \xi(t_1), \ldots \xi(t_n).$$

forms a Markov chain, the problem is posed as follows: Let E denote a finite or countable set. Suppose that a bounded function and a Markov chain with transitional probabilities are defined on E.

(1) Calculate the quantity

$$v(\xi) = \sup_{\tau} M_\xi f(\xi(\tau))$$

where τ ranges over all possible Markov moments.

(2) Find a Markov moment τ_0 for which /128

$$M_\xi f(\xi(\tau_0)) = v(\xi)$$

Here, $v(\xi)$ is called the cost of the game and the Markov moment τ_0 is called the optimal strategy.

This cost of the game $v(\xi)$ is the smallest convex function that is at least equal to the gain function $f(\xi)$.

The solution of the problem posed in the case of countable set E leads to a consideration of an "ε-support" set

$$\mathfrak{M}_\varepsilon = \{\xi : v(\xi) - f(\xi) \leq \varepsilon\}$$

If τ_ε denotes the instant at which x first gets into M_ε, then, for every $\varepsilon > 0$,

$$M_\xi f(\xi(\tau_\varepsilon)) \geq v(\xi) - \varepsilon$$

that is, the "ε-support" sets enable us to find the strategies τ_ε ensuring a gain arbitrarily close to the cost of the game [1].

REFERENCES /129

1. Dynkin, Ye. B., and A.A. Yushkevich, Teoremy i zadachi of protsessakh Markova (Theorems and problems on Markov processes), Fizmatgiz, 1967.
2. Prokhorov, Yu. V., and Yu. A. Rozanov, Teoriya veroyatnostey (probability theory), Fizmatgiz, 1967.

Translated for the National Aeronautics and Space Administration by Scripta Technica, Inc. NASW-2036.